

Rethinking the Role of Evaluation

Finding 6: Rigid and ineffective evaluation practices inappropriately drive programs and hinder the development of effective and efficient prevention initiatives.

The State's requirements for evaluation of prevention programs vary widely from program to program and there are no guidelines to help policy-makers know what is appropriate for a particular initiative.

A recent focus on outcomes has resulted in rigorous program evaluation requirements that often are not aligned with the complexity of violence prevention. Rather than evaluations tailored to the characteristics of a prevention initiative, one standard is applied. As a result, scarce evaluation dollars are spread across too many programs and evaluations fail to provide policy-makers or program managers with useful information.

The State should create a tiered evaluation strategy that would rigorously test new and unproven programs, relax evaluation requirements for proven strategies and develop improved methodologies to evaluate complex strategies and measure community change.

Current Policy

The State has struggled for years with how to evaluate the effectiveness of the programs it funds. Prevention is no exception, and in many ways has proven even more difficult. As a result, evaluation requirements for state-funded prevention programs differ widely. Some programs require rigorous experimental research designs at each site and an independent, overall statewide evaluation. Some only require programs to provide the State with self-evaluations. For some programs, there is no evaluation component.

These variations are not intentional. There are no well established guidelines to help lawmakers establish in law appropriate evaluation requirements for new programs. Often, the evaluation components of new legislative initiatives are drafted by staff who are not trained in research and evaluation, and without input from key stakeholders such as the administering agency and local service providers. Decisions are driven by the availability of funds and current biases regarding evaluation.

This assessment is true for many social service programs. But it is particularly true for programs that are trying to prevent negative behavior or intervene once it has surfaced.

Some policy-makers are skeptical of prevention in general, or believe the State is not investing in the “right” prevention programs. Policy-makers also are increasingly focusing on outcomes rather than inputs.

The State often expects prevention programs to prove that they reduce youth violence – and that they are more cost-effective than other public safety approaches like intervention, treatment and suppression. But a similar burden is not applied to the juvenile justice system, whose primary goal is to rehabilitate young offenders. A study of recidivism among wards released to parole over a 10-month period showed that 59.4 percent were arrested within 24 months of release. Three-and-a-half years later, 76.2 percent had been arrested.⁷⁵ But the State does not tie continued funding for these programs to positive outcomes.

The recent \$120 million annual state allocation for prevention programs requires counties to measure specific outcomes including arrest rates, rates of probation and community service completion. It also requires counties to evaluate programs using true experimental research designs. Most of the 47 Challenge Grant Programs administered by the Board of Corrections also employ rigorous scientific research designs.

Problems with the State’s Evaluation Policies

Policy-makers and the public want to know whether resources invested in youth crime and violence prevention programs are achieving the desired results and whether they are cost effective when compared to the alternatives. But despite their costs, evaluations seldom provide this information. The Commission has identified five problems with the State’s evaluation policies that contribute to these shortcomings.

1. As conducted, evaluations are often not useful to policy-makers. The State often expects evaluations to provide convincing evidence of the efficacy of a particular youth violence prevention program. But young people have multiple influences in their lives – and may be receiving multiple interventions – that could contribute to change. Isolating the effects and attributing change to one program is difficult, if not impossible.

As described throughout this report, research and the experiences of communities have shown that the most effective youth violence prevention strategies are multi-disciplinary and community wide, with multiple factors attempting to positively influence children. Experienced researchers said that these evaluations are complicated and particularly difficult to conduct.

In an attempt to establish cause and effect relationships, researchers employ methodologies that are designed for narrowly defined programs to assess complex strategies. These methods, however, are ill-suited to capture less tangible community indicators of change, like reduced fear of crime or the belief by youth that they are cared for and connected to their community. These and other important measures of success are not measured by traditional scientific evaluations.

As a result, evaluations fail to answer key policy questions about the outcomes, cost-effectiveness and accountability of programs that are needed by policy-makers. Lacking reliable information, policy-makers often disagree about spending more money on prevention – or about how to spend additional resources.

2. Scientific evaluations are expensive. Analysts told the Commission that the cost of scientific evaluations can be as much as the intervention itself, and still not provide information useful to policy-makers and program managers.

Several large foundations that fund youth violence prevention programs have concluded that prevention efforts can be better understood – at far less cost – by using other measures of effectiveness.

To assess the impact of the first half of its 10-year \$60 million Youth Violence Prevention Initiative, the Wellness Foundation awarded \$6 million to the RAND Corporation and the Stanford Center for Research in Disease Prevention. The foundation wanted to assess the effectiveness of interventions at the community level as rigorously and objectively as

Guidelines for Legislative Language for State Program Evaluation

Recognizing that policy-makers often do not receive the guidance they need from program evaluations, the Senate contracted with the California State University to develop guidelines that could be used in drafting evaluation language for new programs. The following questions were intended to help lawmakers decide when and how evaluations should be required:

- Is evaluation of this program an important investment of state resources?
- What questions does the Legislature need to have answered about this program?
- What will it take to answer the Legislature's questions – and can adequate resources be provided?
- What will it take to ensure credible evaluation findings?
- Who should be involved in this evaluation – from inception to results?
- When should evaluation findings be expected from this program?
- What is the role of state agencies in this evaluation?
- What information needs to be available for statewide evaluation?

Source: Dowell, David. 1998. Guidelines for Legislative Language for State Program Evaluation, Faculty Fellows Program, Center for California Studies, California State University.

possible. In the end, the evaluation pointed to important accomplishments such as generating new research, shaping policy-making, and training violence prevention leaders statewide. It did not, however, provide evidence of effectiveness in reducing youth violence as the foundation had hoped.

The foundation concluded: “Despite hopes or expectations, there is no irrefutable empirical proof of causal connections linking changes in violence rates to the foundation’s violence prevention grants.”

Lessons learned from evaluation of the first five years of its initiative are guiding the foundation’s approach to evaluation of the final five years. For one, it will commit far less funding to this second phase of evaluation: just \$1.3 million for the final five years, compared to the \$6 million invested in the first phase. It will diversify its evaluation approaches, de-emphasize academic attribution and focus more on a qualitative analysis – “telling the stories behind the statistics.” Also, the foundation will commit no more than 5 percent of grant-making dollars to evaluation in the future.

Similarly, the State is grappling with how to improve the usefulness of evaluations and lower the cost. A Department of Education task force study concluded that a simple survey for a few sites typically costs \$150,000 annually. More sophisticated evaluations that collect and analyze qualitative and quantitative data from a representative sample of sites cost between \$500,000 to \$1 million annually.⁷⁶

Lessons Learned

The California Wellness Foundation learned important lessons about evaluation.

- *Ask fewer evaluation questions:* Trying to answer too many questions undermined the Foundation’s evaluation from the start. Evaluations should focus on key issues to provide depth rather than breadth.
- *Diversify evaluation approaches:* The impact of complex grantmaking initiatives cannot be assessed by investing only in a traditional, rigorously “scientific” evaluation, especially when measuring changes at the community level.
- *De-emphasize academic attribution:* Despite initial hopes or expectations, there is no irrefutable empirical proof of causal connections linking changes in violence rates to the Foundation’s violence prevention grants.
- *Tell the stories behind the statistics:* The foundation did not invest as much in qualitative analysis of the Initiative (i.e., case studies, individual profiles), and that has resulted in the absence of compelling human stories behind the numbers.
- *Collect lessons learned:* Sharing what was learned (not just what has worked) will be an important contribution to the fields of philanthropy, public health and violence prevention.
- *Link evaluation with broader dissemination:* What to evaluate must be tied to an overall strategy about why, how and to whom we communicate evaluation findings.

3. Evaluations are seldom useful to program managers. The evaluation needs of local program managers are different than those of the State. The State designs evaluations to show whether a program is effective in achieving articulated goals. Those results often are not known until a program has been in effect for several years. Program managers need evaluations to provide ongoing feedback so they can make adjustments and improvements to their programs on a day-to-day basis. Information about what worked and what didn't when their program is nearing the end of its grant cycle is of little use.

Evaluation requirements can actually compromise efforts to help children. Sometimes useful information is kept from managers to ensure the integrity of evaluations. Sometimes the needs of researchers to publish certain kinds of evaluations drive the design of evaluation – rather than the needs of policy-makers or program managers to improve their performance.

The Commission also heard concerns from program managers that researchers often do not reach out to the community, reflect the community, or meet the needs of the community. The director of a community-wide coalition described evaluators as “the enemy who comes into our community and studies us.”⁷⁷ She said to be effective, researchers need to see their role as partners and friends of the communities they study.

Positive relationships between communities and program evaluators foster more effective use of available data and evaluation by communities. A researcher with the California Children and Families First Program said that local Proposition 10 commissions are encouraged to set aside 10 percent of their funds for evaluation. He confided that the underlying intent of that guideline is to get local programs interested in collecting meaningful data, a practice the state commission hopes will continue even if funding ends.⁷⁸

4. Children who could benefit are denied services. In some cases, youth who could benefit from services are denied them because of evaluation requirements that insist on the use of comparison groups. In the 47 Challenge Grant Programs, 10,420 youth have been assigned to comparison groups. Those youth receive standard probation services, rather than the enhanced services offered the treatment group.⁷⁹ A Fresno County probation officer lamented turning youth away who wanted to be tutored and mentored in order to meet the evaluation requirements for a state-funded program.

In some cases it may be necessary to deny services to some youth to test a truly experimental idea. But the State should not support evaluation that denies services to children that some research and experience show are

effective, like after-school programs, tutoring, mentoring and Boys & Girls Clubs.

For decades parents who could afford these enhancements, purchased them for their children knowing – absent scientific research – that they were beneficial. But current policy requires willing children in troubled families and neighborhoods to be turned away from state-funded opportunities for the sake of an evaluation.

5. *Community practitioners lack evaluation skills.* The Commission was told that program managers at the community level often lack the expertise necessary to conduct useful evaluations. A member of the Commission's advisory committee said: "Local folks are not used to documenting outcomes. They are too busy training and teaching. The whole idea of

Evaluation in Illinois

In Illinois, a collaborative effort by the Illinois Center for Violence Prevention and the Illinois Violence Prevention Authority has created an Evaluation Resource Institute.

Its purpose is to provide communities with the tools they need to evaluate their programs effectively and to disseminate the results statewide. In addition, the institute offers training on evaluation issues.

evaluation needs the coaching element. Consolidate programs so there aren't so many small grants, each with their own data collection or evaluation requirements. Prioritize and build in local coaches that will help people learn how to document outcomes."⁸⁰ A program manager of a community-based organization said the State should assist communities with evaluation, provide training and support local agencies when they make mistakes.⁸¹

The Superintendent of Public Instruction told the Commission that a close look at state-funded grant programs shows that for some programs, the administering agency is provided resources to provide local agencies with technical assistance. For other programs, resources are not provided.⁸²

6. *Too little time is provided.* As described in Finding 3, most state grant programs are for periods of one to three years. The State Department of Education told the Commission that new initiatives often do not achieve full operation until the third year of funding, making meaningful evaluation of long-term outcomes often impractical. Lack of time to demonstrate results often means that decisions to continue – or discontinue – funding are made in the absence of adequate information. The Department of Finance told the Commission that because of inadequate evaluation, fiscal decisions are routinely made in the absence of knowledge about what works.

7. *Evaluations are not strategically coordinated.* Evaluation requirements and expectations are not coordinated among state agencies that administer prevention programs or among those that administer similar programs, like gang violence prevention for example. Moreover, evaluation requirements among programs administered within an agency administering

multiple prevention programs are not even coordinated. Absent coordination and conformity to an accepted standard, there is no easy way to compare particular evaluations. In other words, it is difficult to compare evaluations of diverse programs or to assess the quality of the evaluations.

Goals for Evaluation

The State needs to align evaluation to its need for information. With the assistance of its advisory committee and other experts the Commission identified the following desired outcomes for evaluation.

- ***Information to help policy-makers determine how much to invest in evaluation and how to fashion that investment.*** Evaluation requirements and resource needs depend on the size, scope, type of project and the kinds of measures required. For example, survey data on a relatively small number of similar sites could be done for a modest cost. Where quantitative and qualitative data are required to provide outcome measures, costs increase. To make good decisions, policy-makers need criteria to determine what would be an appropriate evaluation for specific prevention initiatives. They need information that will permit them to tailor evaluations to specific policy interventions, rather than applying one standard to all programs. Costs and time required to achieve objectives must be taken into account.
- ***To know if prevention resources are being spent wisely.*** Policy-makers and the public want to know if prevention resources are being spent wisely. Policy-makers need to know whether strategies in place prevent youth crime and violence and identify problems with those that are not working. And they need to know whether strategies are cost-effective.
- ***The ability to understand community change.*** Evaluation methodologies should be developed that can measure community changes related to violence. Diverse evaluation methodologies that include qualitative and quantitative measures of success should be developed.
- ***Information that can guide program management.*** Program managers need continuous feedback to identify and respond to problems as they arise. To help prevention practitioners effectively manage programs, management evaluation tools and indicators of community health should be developed.
- ***Provide promising practices information.*** Evaluation should allow for the effective documentation of promising and proven practices and guide faithful program replication.

Devising Effective Evaluation Policy

The State should rethink the evaluation requirements imposed on prevention programs. By aligning its policies with the reality of complex violence prevention strategies, scarce evaluation dollars could be expended effectively to provide policy-makers, the public and program managers with useful information. The alternatives devised by other entities could be instructive to the State.

“Earmark 10 percent and Centralize Evaluation”

In 1996 Congress directed the U.S. Attorney General to evaluate the effectiveness of over \$3 billion in federal crime prevention grants. Specifically, Congress required that the evaluation focus on the effectiveness of programs designed to prevent youth violence. Evaluators concluded that the quality and quantity of program evaluations are not adequate to guide national efforts to reduce serious crime. It recommended that the statutory evaluation plan of the Department of Justice be reformed to provide the scientific tools necessary for effective evaluation.

The report recommended that Congress earmark 10 percent of all funding for local crime prevention efforts to a central evaluation office in the Department of Justice. The central evaluation office would distribute those funds for rigorous scientific impact evaluations that could be generalized to other locations. Those funds would add to the total funding for which a local grantee is eligible – serving as an incentive for cooperation with the evaluation plan.

Recognizing the expense of rigorous scientific evaluation, the report recommended that an additional 10 percent of all funding for local crime prevention be set aside for evaluations to be conducted by the central evaluation office.

“Measure Success Many Ways”

The California Wellness Foundation concluded that there are a variety of ways to gauge the success of prevention and prevention sponsors should invest in different evaluation approaches using diverse measures of success.⁸³

The David and Lucille Packard Foundation and the California Endowment repeated that they often find more value in intangible characteristics of successful efforts and recommended that the State rethink the evidence required as proof of effectiveness. The Packard Foundation said that while it is important to know if programs work, “it may be that one of the most

critical results of just trying to do something about youth violence has been to create a greater sense of community mission, cohesiveness and connectedness, and a sense among youth that they are cared for and valued.”⁸⁴

“Coordinate and Collaborate”

In its 1995 report the Attorney General’s Policy Council on Violence Prevention recommended that public and private organizations coordinate investments in research, assessing which policies and programs effectively prevent violence. It said that all violence prevention efforts should include an impact-evaluation component, a cost-effectiveness component and a “learn-as-we-go” approach that allows for improving programs based on evaluation data. “By building in provisions to learn throughout the process and by utilizing evaluation data, violence prevention and intervention programs can be improved along the way.”⁸⁵

The Policy Council also recommended that California institutions of higher learning develop and implement programs to train researchers in violence prevention and research.

Summary: A “Tiered” Approach to Evaluation

Without a solid evaluation policy, the State’s desire to measure outcomes has placed rigid and often inappropriate evaluation requirements on many prevention programs. As a result, scarce evaluation dollars are spread across too many programs, key policy questions are not answered, and program managers cannot use data to improve services.

A strategic, tiered approach to evaluation could provide policy-makers, the public and practitioners with the information they need and would be more cost effective. Such an approach could require rigorous evaluation of selected, untested strategies that represent a significant public investment. For strategies that have reliable evidence of success – as the result of scientific research or the collective experiences of practitioners and participants – evaluation could measure faithful replication, effective management and fiscal accountability.

Recommendation 6: To inform policy-makers, practitioners and the public, the State should adopt a strategy for evaluating prevention efforts. Specifically, the Youth Violence Prevention Institute should:

- ❑ ***Develop and recommend effective evaluation methods.*** The plan should distinguish between the level of evaluations that are needed to test experimental strategies, versus those that can determine if proven programs are being faithfully replicated. Experimental programs – particularly those that represent significant public expenditures – should be rigorously evaluated. Proven programs should only be evaluated for fiscal accountability, program implementation and management effectiveness.
- ❑ ***Help develop community indicators.*** Prevention providers need to be accountable to their communities and the State for improving the lives of young people. The institute should work with the Youth Violence Prevention Coordinating Council to develop indicators of community health that will assist communities to identify problems and measure progress.
- ❑ ***Provide a way to understand community change.*** The evaluation strategy should advance methodologies to assess complex efforts and effectively measure community change, based on goals and indicators of community health. The strategy should include exploration of more efficient ways to conduct evaluations.
- ❑ ***Develop evaluation tools for program management.*** Local service providers need and want to use evaluation to assess progress and improve services. The Institute should develop tools to help communities evaluate and improve program management.